

# Open Research Online

---

The Open University's repository of research publications  
and other research outputs

## Towards a fuzzy domain ontology extraction method for adaptive e-learning

### Journal Item

#### How to cite:

Lau, R. Y. K.; Song, Dawei; Li, Yuefeng; Cheung, T. C. H. and Hao, Jin-Xing (2009). Towards a fuzzy domain ontology extraction method for adaptive e-learning. IEEE Transactions on Knowledge and Data Engineering, 21(6) pp. 800–813.

For guidance on citations see [FAQs](#).

© 2009 IEEE

Version: [\[not recorded\]](#)

Link(s) to article on publisher's website:  
<http://dx.doi.org/doi:10.1109/TKDE.2008.137>

---

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data [policy](#) on reuse of materials please consult the policies page.

---

[oro.open.ac.uk](http://oro.open.ac.uk)

# Toward a Fuzzy Domain Ontology Extraction Method for Adaptive e-Learning

Raymond Y.K. Lau, *Senior Member, IEEE*, Dawei Song, Yuefeng Li, *Member, IEEE*,  
Terence C.H. Cheung, *Member, IEEE*, and Jin-Xing Hao

**Abstract**—With the widespread applications of electronic learning (e-Learning) technologies to education at all levels, increasing number of online educational resources and messages are generated from the corresponding e-Learning environments. Nevertheless, it is quite difficult, if not totally impossible, for instructors to read through and analyze the online messages to predict the progress of their students on the fly. The main contribution of this paper is the illustration of a novel concept map generation mechanism which is underpinned by a fuzzy domain ontology extraction algorithm. The proposed mechanism can automatically construct concept maps based on the messages posted to online discussion forums. By browsing the concept maps, instructors can quickly identify the progress of their students and adjust the pedagogical sequence on the fly. Our initial experimental results reveal that the accuracy and the quality of the automatically generated concept maps are promising. Our research work opens the door to the development and application of intelligent software tools to enhance e-Learning.

**Index Terms**—Domain ontology, ontology extraction, text mining, fuzzy sets, concept map, e-Learning.

## 1 INTRODUCTION

ELECTRONIC learning (e-Learning) refers to the application of information and communication technologies (e.g., Internet and multimedia) to enhance ordinary classroom teaching and learning [3], [46]. With the maturity of the technologies such as the Internet and the decreasing cost of the hardware platforms, more institutions are adopting e-Learning as a supplement to traditional instructional methods [44], [46]. One of the potential advantages of modern e-Learning technologies is that they can facilitate *adaptive learning* such that instructors can dynamically revise and deliver instructional materials in accordance with the learners' current progress. In general, adaptive teaching and learning refers to the use of what is known about learners, a priori or through interactions, to alter how a learning experience unfolds, with the aim of improving learners' success and satisfaction [16]. The current state-of-the-art of e-Learning technologies support automatic collection of learners' performance data (e.g., via online quiz) [11]. However, few of the existing e-Learning technologies can support automatic analysis of learners' progress in terms of the knowledge structures the learners have acquired. In this paper, we illustrate a methodology of automatically constructing concept maps to characterize

learners' understanding for a particular topic; thereby instructors can conduct adaptive teaching and learning based on the learners acquired knowledge structures as reflected on the concept maps. In particular, our concept map generation mechanism is underpinned by a context-sensitive text mining method [21] and a novel fuzzy domain ontology extraction algorithm.

It is generally accepted that ontology refers to a formal specification of conceptualization [13]. Ontology can take the simple form of a taxonomy of concepts (i.e., lightweight ontology), or the more comprehensive representation of comprising a taxonomy as well as the axioms and constraints which characterize some prominent features of the real world (i.e., heavy weight ontology) [4]. Domain ontology is one kind of ontology which is used to represent the knowledge for a particular type of application domain [8]. On the other hand, concept maps are used to elicit and represent the knowledge structure such as concepts and propositions as perceived by individuals [33]. Concept maps are similar to ontology in the sense that both of these tools are used to represent concepts and the semantic relationships among concepts. However, ontology is a formal knowledge representation method to facilitate human and computer interactions and it can be expressed by using formal semantic markup languages such as RDF and OWL [15], whereas concept map is an informal tool for humans to specify semantic knowledge structure. We only focus on the automatic extraction of lightweight domain ontology in this paper. More specifically, the lightweight fuzzy domain ontology is used to generate concept maps to represent learners' knowledge structures.

The main contributions of our research work are twofold: from the theoretical stand point, we contribute to the development of a novel fuzzy domain ontology extraction method which alleviates the knowledge acquisition bottleneck of manually constructing domain ontologies. Since ontology extraction from text often involves uncertainty

- R.Y.K. Lau, T.C.H. Cheung, and J.-X. Hao are with the Department of Information Systems, City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong. E-mail: {raylau, is.tc, jinxihao}@cityu.edu.hk.
- D. Song is with the Knowledge Media Institute, The Open University, Walton Hall, MK7 6AA Milton Keynes, UK. E-mail: d.song@open.ac.uk.
- Y. Li is with the Faculty of Information Technology, Queensland University of Technology, GPO Box 2434, Brisbane QLD 4001, Australia. E-mail: y2.li@qut.edu.au.

Manuscript received 21 Jan. 2008; accepted 16 June 2008; published online 25 June 2008.

Recommended for acceptance by Q. Li, R.W.H. Lau, D. McLeod, and J. Liu. For information on obtaining reprints of this article, please send e-mail to: tkde@computer.org, and reference IEEECS Log Number TKDE-2008-01-0046. Digital Object Identifier no. 10.1109/TKDE.2008.137.

(e.g., which messages (objects) are associated with which concepts (classes)), an uncertainty representation and management mechanism is required to address such an issue. It is believed that the notions of fuzzy set and fuzzy relation provide a sound and rigorous method to represent knowledge with uncertainty [54]. One of our contributions is the development of a formal fuzzy domain ontology model which is underpinned by fuzzy sets and fuzzy relations. Moreover, based on the concept of subsumption, we have developed a fuzzy domain ontology extraction algorithm for the automatic extraction of domain ontologies from text. From the practical standpoint, our research work opens the door to the development of intelligent software tools for enhancing e-Learning technologies. In particular, we have demonstrated how to apply the context-sensitive text mining method and the fuzzy domain ontology extraction algorithm to automatically generate concept maps to reveal the knowledge structures of students who are engaging in e-Learning. As a result, instructors can adjust the pedagogy with respect to the student's current progress on the fly. This is the so-called adaptive teaching and learning approach [16].

The remainder of this paper is organized as follows: A framework of fuzzy domain ontology-based concept map generation is highlighted in Section 2. The cognitive and linguistic foundations of the proposed context-sensitive text mining method for concept extraction are illustrated in Section 3. Then, the computational algorithm of the fuzzy domain ontology extraction method is depicted in Section 4. Section 5 explains how the fuzzy domain ontology extraction method is applied to adaptive e-Learning. Section 6 describes the evaluation of the proposed fuzzy domain ontology-based concept map generation mechanism. Section 7 highlights previous research in the related area and compares these research works with ours. Finally, we offer concluding remarks and describe future direction of our research work.

## 2 A FRAMEWORK FOR AUTOMATIC CONCEPT MAP GENERATION

It has been pointed out that the main challenge of automatic ontology extraction from textual databases is the removal of noisy concepts and relations [27], [28]. Based on this observation, our domain ontology extraction methodology in general and the concept map generation process in particular are designed to effectively filter the nonrelevant concepts and concept relations from the concept space. Fig. 1 depicts the proposed methodology of automatically generating concept maps from a collection of online messages posted to blogs, emails, chat rooms, Web pages, and so on. The collection of messages is treated as a textual corpus. At the document parsing stage, our document parser will scan each message to analyze the lexico-syntactic elements embedded in the message. For instance, stop words such as *a*, *an*, *the*, and so on are removed from the message since these words appear in any contexts and they cannot provide useful information to describe a domain concept. For our implementation, a stop word file is constructed based on the standard stop word file used in the SMART retrieval system [38]. Different customizations are required for processing different kinds of documents.

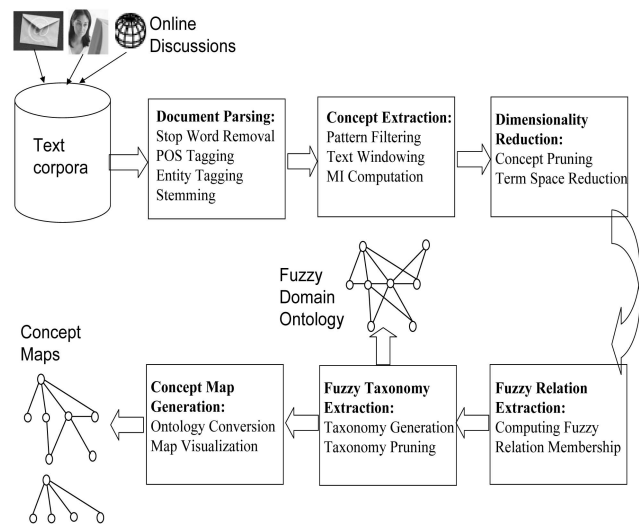


Fig. 1. A framework for automatic concept map generation.

For example, we need to extend the SMART stop word file by including stop words such as “home,” “contact,” “Web,” and “site” for parsing Web pages.

Lexical patterns are identified by applying Part-of-Speech (POS) tagging to the source documents. We develop our POS tagger based on the WordNet API.<sup>1</sup> For named-entity detection (e.g., people's names and organizations' names), we employ BBN's *IdentiFinder* [1]. However, for the e-Learning application reported in this paper, we do not make use of the entity tags for concept extraction. We simply treat each named-entity as a noun for subsequent linguistic pattern mining. After the tagging process, each token is stemmed according to the Porter stemming algorithm [36]. During the concept extraction stage (Section 4.1), certain linguistic patterns are ignored to reduce the generation of noisy concepts. For example, ontology engineers or instructors in the case of e-Learning applications will specify the mining focus on certain linguistic patterns such as “Noun Noun,” “Adjective Noun,” and “Verb Noun.” The text mining program will then focus on finding the term association information and collect the statistical data for those patterns only. Not only does it reduce the generation of noisy concepts but also improve the computational efficiency of our ontology extraction process. A text windowing process will be conducted by scanning adjacent tokens within a predefined window size of 5 to 10 words from left to right over all the documents. At the end of the windowing process, an information theoretic measure is applied to compute the co-occurrence statistics between the targeting linguistic patterns and other tokens appearing in the same text window across the corpus. Thereby, context vectors [17], [41] can be created to describe the semantic of the extracted concepts.

In addition, to filter out nonrelevant domain concepts, the occurrence of a concept across different domains (e.g., corpora) will be assessed (Section 4.2). The basic intuition is that a concept frequently appears in a specific domain (corpus) rather than many different domains is more likely

1. <http://wordnet.princeton.edu/>.

to be a relevant domain concept. Those concepts with relevance scores below a certain threshold will not be used for taxonomy generation. To produce accurate concept representations, a dimensionality reduction method is applied to the filtered concept space to minimize the terms (features) used to characterize the concepts based on the principle of minimal information loss (Section 4.3). After concept space reduction, the subsumption relationships among the domain concepts are computed according to our fuzzy relation membership function (Section 4.4). A taxonomy of fuzzy domain concepts is then constructed according to our fuzzy domain ontology extraction algorithm (Section 4.5). Finally, our visualization mechanism converts the fuzzy domain ontology to concept maps and displays them on our Web-based e-Learning platform.

Before illustrating the computational details of our fuzzy domain ontology extraction method in the remaining sections, we should give a precise definition of what we mean by lightweight fuzzy domain ontology. Our proposed model of fuzzy domain ontology is underpinned by fuzzy sets and fuzzy relations [54].

**Definition 1 (Fuzzy set).** A fuzzy set  $\mathcal{F}$  consists of a set of objects drawn from a domain  $X$  and the membership of each object  $x_i$  in  $\mathcal{F}$  is defined by a membership function  $\mu_{\mathcal{F}} : X \mapsto [0, 1]$ . For the special case of a crisp set, the crisp membership function has the mapping  $\mu_{\mathcal{F}} : X \mapsto \{0, 1\}$ .

**Definition 2 (Fuzzy relation).** A fuzzy relation  $R_{XY}$  is defined as the fuzzy set  $\mathcal{R}$  on a domain  $X \times Y$ , where  $X$  and  $Y$  are two crisp sets. The membership of each object  $(x_i, y_i)$  in  $\mathcal{R}$  is defined by a membership function  $\mu_{\mathcal{R}} : X \times Y \mapsto [0, 1]$ .

**Definition 3 (Fuzzy ontology).** A fuzzy ontology is a 6-tuple  $\text{Ont} = \langle X, A, C, R_{XC}, R_{AC}, R_{CC} \rangle$ , where  $X$  is a set of objects,  $A$  is the set of attributes describing the objects, and  $C$  is a set of concepts (classes). The fuzzy relation  $R_{XC} : X \times C \mapsto [0, 1]$  assigns a membership to the pair  $(x_i, c_i)$  for all  $x_i \in X, c_i \in C$ , the fuzzy relation  $R_{AC} : A \times C \mapsto [0, 1]$  defines the mapping from the set of attributes  $A$  to the set of concepts  $C$ , and the fuzzy relation  $R_{CC} : C \times C \mapsto [0, 1]$  defines the strength of the subclass/superclass relationships among the set of concepts  $C$ .

Fig. 2 illustrates an example of our lightweight fuzzy domain ontology with reference to the above definitions. In this example,  $X = \{x_1, \dots, x_7\}$ ,  $A = \{a_1, \dots, a_6\}$ , and  $C = \{c_1, \dots, c_5\}$  are assumed. The fuzzy relation  $R_{CC}$  (i.e., subclass/superclass relationships) is denoted as  $R_{CC}(c_x, c_y)$  and two examples, e.g.,  $R_{CC}(c_2, c_1) = 0.5$  and  $R_{CC}(c_3, c_1) = 0.4$ , are shown in Fig. 2. The  $R_{XC}$  relation describes the membership of an object for a particular class (concept). For instance, object  $x_7$  is considered belonging to the class  $c_4$  with a membership value of 0.3. For the e-Learning application, the ontology can represent which online message (i.e., an object) created by a learner is associated with certain concepts to facilitate the analysis of students' learning states. To improve the readability of Fig. 2, the partial associations between concepts and attributes (i.e.,  $R_{AC}$ ) are not depicted. For a concept such as "commercial bank," we may find a property term (i.e., attribute) "customer" describing the concept. However, the term "customer" may also be used to describe other

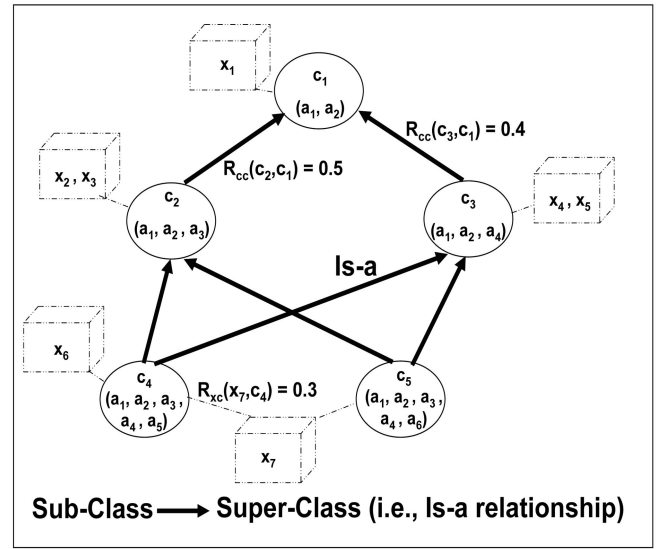


Fig. 2. An example of fuzzy domain ontology.

concepts such as "book shop" to a certain degree. Our lightweight fuzzy domain ontology model is able to represent the partial associations among concepts and the underlying property terms. Based on the idea of formal concept analysis [7],  $X$  is the *extent* of the concepts  $C$ , and  $A$  is the *intent* which defines the properties of  $C$ . According to the idea of subsumption, the subconcept/superconcept relation ( $R_{CC}$ ) can be defined by:

**Definition 4 (Fuzzy subsumption).** With respect to an arbitrary  $\alpha$ -cut level, a concept  $c_x \in C$  is the subconcept of another superconcept  $c_y \in C$  if and only if  $\forall a_i \in \{z \in A | \mu_{R_{AC}}(z, c_y) \geq \alpha\}, \mu_{R_{AC}}(a_i, c_x) \geq \alpha$ . Alternatively, from an extensional perspective, a concept  $c_x \in C$  is the subconcept of another superconcept  $c_y \in C$  if and only if  $\forall x_i \in \{z \in X | \mu_{R_{XC}}(z, c_x) \geq \alpha\}, \mu_{R_{XC}}(x_i, c_y) \geq \alpha$  with respect to an arbitrary  $\alpha$ -cut level.

Definition 4 can be explained as follows: If the membership of every attribute  $a_i \in A$  for the concept  $c_y \in C$  is greater than or equal to a certain threshold  $\alpha$ , the membership of the corresponding attribute  $a_i$  for the concept  $c_x \in C$  is also greater than or equal to  $\alpha$ , then the concept  $c_x$  is the subconcept of  $c_y$ . As can be seen, the crisp subsumption relation is only a special case of the generalized fuzzy subsumption relation in that the threshold value  $\alpha = 1$  is established for the crisp case. In other words, if it is true that every attribute  $a_i \in A$  characterizing the concept  $c_y$  implies that it also characterizes the concept  $c_x$ , the concept  $c_x$  is the subconcept of  $c_y$ .

### 3 THE COGNITIVE AND LINGUISTIC FOUNDATIONS

From a human cognitive perspective, humans learn a new concept by associating the contexts (experiential background) in which the concept appears [12]. Our concept extraction method is motivated based on the above observation. In particular, our context-sensitive text mining approach is developed according to the *distributional hypothesis* which assumes that terms (concepts) are similar

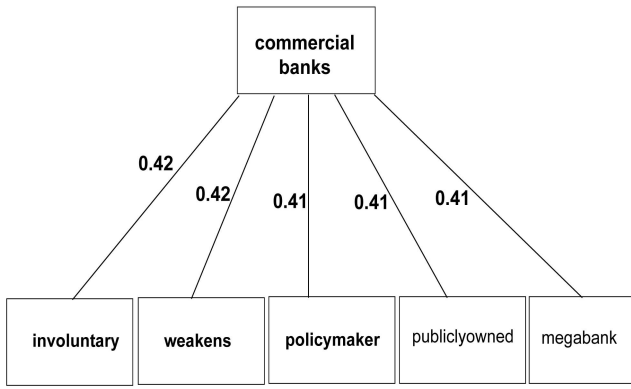


Fig. 3. Domain-specific semantics of the concept “Commercial Banks.”

according to the extent that they share similar linguistic contexts [14]. The notion of *collocational expressions* from computational linguistic is applied to extract the semantics of certain lexical elements (e.g., concepts) from text corpora [42]. Collocational expressions are groups of words related in meaning, and the constituent words of an expression are frequently found in a near loci of a few adjacent words in a textual unit [42], [45]. Collocational expressions provide the contexts to extract the semantics of concepts embedded in natural language texts such as net news, blogs, emails, or Web documents [35]. In computational linguistic, a term refers to one or more tokens (words), and a term could also been taken as a concept if it carries recognizable meaning with respect to a context (domain) [17], [31].

In the field of information retrieval (IR), the notion of *context vectors* [17], [41] has been proposed to construct computer-based representations of concepts (i.e., linguistic class). In this approach, a concept is represented by a vector of words (property terms) and their numerical weights. The weight of a word indicates the extent to which the particular word is *associated* with the underlying concept. Fig. 3 shows that the concept “commercial banks” is represented by the property terms such as involuntary, weakens, policymaker, publiclyowned, and megabank. Indeed, this is an interesting example generated from the Reuters-21578 corpus.<sup>2</sup> The context vector of “commercial banks” is shown as follows:

Concept: commercial banks

Context Vector:

$\langle (involuntary, 0.42), (weakens, 0.42), (policymaker, 0.41), (publiclyowned, 0.41), (megabank, 0.41) \rangle$ .

A context vector can be seen as a point in a multi-dimensional geometric information space with each dimension representing a property term. It should be noted that the meanings (senses) of “commercial banks” is “a financial institution that accepts deposits and makes loans and provides other services for the public” as defined in WordNet [30], which is quite different from that discovered by our context-sensitive text mining method [21] as shown in Fig. 3. Static lexicons such as WordNet can only capture the

general lexical knowledge of a concept, but fails to represent domain-specific information such as the US financial market as described by the Reuters-21578 collection. A linguistic concept such as “commercial banks” can be taken as a class (set) with respect to the fuzzy set framework. A property term such as “publiclyowned” can then be treated as an attribute describing the concept to a certain degree (i.e.,  $\mu_{RAC}(publiclyowned, commercialbanks) = 0.41$ ).

## 4 AUTOMATIC FUZZY DOMAIN ONTOLOGY EXTRACTION

### 4.1 Concept Extraction

Our text mining method is specifically designed to filter noisy concepts. After standard document preprocessing such as stop word removal, POS tagging, and word stemming [39], a *windowing process* is conducted over the collection of documents. The windowing process can help reduce the number of noisy terms. For each document (e.g., Net news, Web page, and email), a *virtual window* of  $\delta$  words is moved from left to right, one word at a time until the end of a textual unit (e.g., a sentence) is reached. Within each window, the statistical information among tokens is collected to develop collocational expressions. Such a windowing process has successfully been applied to text mining before [21], [35]. The windowing process is repeated for each document until the entire collection has been processed. According to previous studies, a text window of 5 to 10 terms is effective [17], [35], and so we adopt this range as the basis to perform our windowing process. To improve computational efficiency and filter noisy concepts, only the specific linguistic patterns (e.g., Noun Noun and Adjective Noun) defined by the user will be analyzed. After parsing the whole corpus, the statistical data (e.g., mutual information) about the potential concepts is collected by our statistical token analyzer. If the association weight between a concept and a term is below a predefined threshold value  $\zeta$ , it will be discarded from the context vector of the concept.

For statistical token analysis, several information theoretic methods are employed. Mutual information has been applied to collocational analysis [35], [47] in previous research. Mutual information is an information theoretic method to compute the dependency between two entities and is defined by [43]

$$MI(t_i, t_j) = \log_2 \frac{Pr(t_i, t_j)}{Pr(t_i)Pr(t_j)}, \quad (1)$$

where  $MI(t_i, t_j)$  is the mutual information between term  $t_i$  and term  $t_j$ .  $Pr(t_i, t_j)$  is the joint probability that both terms appear in a text window, and  $Pr(t_i)$  is the probability that a term  $t_i$  appears in a text window. The probability  $Pr(t_i)$  is estimated based on  $\frac{|w_t|}{|w|}$ , where  $|w_t|$  is the number of windows containing the term  $t$  and  $|w|$  is the total number of windows constructed from a corpus. Similarly,  $Pr(t_i, t_j)$  is the fraction of the number of windows containing both terms out of the total number of windows.

2. <http://www.daviddlewis.com/resources/testcollections/>.

We develop *Balanced Mutual Information* (BMI) to compute the degree of association among tokens. This method considers both term presence and term absence as the evidence of the implicit term relationships:

$$\begin{aligned} \mu_{c_i}(t_j) &\approx BMI(t_i, t_j), \\ &= \beta \times \left[ Pr(t_i, t_j) \log_2 \left( \frac{Pr(t_i, t_j) + 1}{Pr(t_i)Pr(t_j)} \right) \right. \\ &\quad \left. + Pr(\neg t_i, \neg t_j) \log_2 \left( \frac{Pr(\neg t_i, \neg t_j) + 1}{Pr(\neg t_i)Pr(\neg t_j)} \right) \right] \\ &\quad - (1 - \beta) \times \left[ Pr(t_i, \neg t_j) \log_2 \left( \frac{Pr(t_i, \neg t_j) + 1}{Pr(t_i)Pr(\neg t_j)} \right) \right. \\ &\quad \left. + Pr(\neg t_i, t_j) \log_2 \left( \frac{Pr(\neg t_i, t_j) + 1}{Pr(\neg t_i)Pr(t_j)} \right) \right], \end{aligned} \quad (2)$$

where  $\mu_{c_i}(t_j)$  is the membership function to estimate the degree of a term  $t_j \in A$  belonging to a concept  $c_i \in C$ .  $\mu_{c_i}(t_j)$  is the computational mechanism for the relation  $R_{AC}$  defined in the fuzzy domain ontology  $Ont = \langle X, A, C, R_{XC}, R_{AC}, R_{CC} \rangle$ . The membership function  $\mu_{c_i}(t_j)$  is indeed approximated by the BMI score. The weight factor  $\beta > 0.5$  is used to control the relative importance of two kinds of evidence (positive and negative).

Other measures that are used to estimate the membership values of  $t_j \in c_i$  include Jaccard (JA), conditional probability (CP), Kullback-Leibler divergence (KL), Expected Cross Entropy (ECH) [19], and Normalized Google Distance (NGD) [6]. For (7), the term  $|w_{c_i}|$  means the number of virtual text windows containing the concept  $c_i$  and the term  $|w_{c_i, t_j}|$  refers to the number of virtual text windows containing both  $c_i$  and  $t_j$ . After computing term-concept association weights using any one of the methods mentioned above, the association weights are subject to linear scaling using  $v_{Norm} = \frac{v - v_{min}}{v_{max} - v_{min}}$ . As a result, all the term-concept association weights fall into the unit interval  $\forall c_i \in C, t_j \in X, \mu_{c_i}(t_j) \in [0, 1]$ . As NGD is a distance measure, we use the dual function to generate the membership values (e.g.,  $\mu_{c_i}(t_j) = 1 - NGD_{Norm}(c_i, (t_j))$ ). A  $\zeta$ -cut is applied to discard terms from the potential concept if their memberships are below the threshold  $\zeta$ . It should be noted that the constituent terms of a concept are always implicitly associated with the concept itself with the maximal membership 1:

$$\begin{aligned} \mu_{c_i}(t_j) &\approx Jacc(c_i, t_j) \\ &= \frac{Pr(c_i \wedge t_j)}{Pr(c_i \vee t_j)}, \end{aligned} \quad (3)$$

$$\begin{aligned} \mu_{c_i}(t_j) &\approx Pr(c_i | t_j) \\ &= \frac{Pr(c_i, t_j)}{Pr(t_j)}, \end{aligned} \quad (4)$$

$$\begin{aligned} \mu_{c_i}(t_j) &\approx KL(c_i || t_j) \\ &= \sum_{c_i \in C} Pr(c_i | t_j) \log_2 \frac{Pr(c_i | t_j)}{Pr(c_i)}, \end{aligned} \quad (5)$$

$$\begin{aligned} \mu_{c_i}(t_j) &\approx ECH(t_j, c_i) \\ &= Pr(t_j) \sum_{c_i \in C} Pr(c_i | t_j) \log_2 \frac{Pr(c_i | t_j)}{Pr(c_i)}, \end{aligned} \quad (6)$$

$$\begin{aligned} \mu_{c_i}(t_j) &\approx NGD(c_i, t_j) \\ &= \frac{\max\{\log_2 |w_{c_i}|, \log_2 |w_{t_j}|\} - \log_2 |w_{c_i, t_j}|}{\log_2 |w| + 1 - \min\{\log_2 |w_{c_i}|, \log_2 |w_{t_j}|\}}. \end{aligned} \quad (7)$$

## 4.2 Concept Pruning

To further filter the noisy concepts, we adopt the TFIDF [39] like heuristic to perform the filtering process. Similar approach has also been used in ontology learning [32]. For example, if a concept is significant for a particular domain, it will appear more frequently in that domain when compared with its appearance in other domains. The following measure is used to compute the relevance score of a concept:

$$Rel(c_i, D_j) = \frac{Dom(c_i, D_j)}{\sum_{k=1}^n Dom(c_i, D_k)}, \quad (8)$$

where  $Rel(c_i, D_j)$  is the relevance score of a concept  $c_i$  in the domain  $D_j$ . The term  $Dom(c_i, D_j)$  is the domain frequency of the concept  $c_i$  (i.e., number of documents containing the concept divided by the total number of documents in the corpus). The higher the value of  $Rel(c_i, D_j)$ , the more relevant the concept is for domain  $D_j$ . Based on empirical testing, we can estimate a threshold  $\varpi$  for a particular domain. Only the concepts with relevance scores greater than the threshold will be selected. For each selected concept, its context vector is also expanded based on the synonymy relation defined in WordNet [30]. This is in fact a *smoothing* procedure [7]. The intuition is that some terms characterizing a particular concept may not co-occur with the concept in a corpus. To make our ontology extraction method more robust, we need to consider these missing properties. For instance, the context vector “commercial banks” of our example will be expanded with the term “deposits” based on the synonymy relation of WordNet, and a default membership will be assigned to such a term.

## 4.3 Dimensionality Reduction

In order to reduce the dimensionality of the concept space, unsupervised mapping techniques such as *Singular Value Decomposition* (SVD) [9], [10] is applied to the *Term-Concept Association Matrix*,  $\mathbf{R}$ , which is formed by the membership values  $\mu_{c_i}(t_j)$  for all term  $t_j \in A$  belonging to some concepts  $c_i \in C$  after the previous concept pruning stage. In general,  $\mathbf{R}$  can be expressed by any rectangular  $m \times p$  matrix, whereas  $m = |A|$  and  $p = |C|$ . The general complexity of computing SVD is  $O(\min(mp^2, pm^2))$ . As the number of concepts has been reduced to  $n$  (where  $n \ll p$ ) by the concept pruning process, the actual computational complexity of our SVD process is reduced to  $O(\min(mn^2, nm^2))$ . By controlling the pruning parameter  $\varpi$ , our SVD can scale

up for a large collection of messages. Each element in  $\mathbf{R}$  represents the membership value  $\mu_{c_i}(t_j)$ , i.e.

$$\mathbf{R} = \begin{pmatrix} \mu_{c_1}(t_1) & \mu_{c_2}(t_1) & \dots & \mu_{c_i}(t_1) & \dots & \mu_{c_n}(t_1) \\ \mu_{c_1}(t_2) & \mu_{c_2}(t_2) & \dots & \mu_{c_i}(t_2) & \dots & \mu_{c_n}(t_2) \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \mu_{c_1}(t_j) & \mu_{c_2}(t_j) & \dots & \mu_{c_i}(t_j) & \dots & \mu_{c_n}(t_j) \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \mu_{c_1}(t_m) & \mu_{c_2}(t_m) & \dots & \mu_{c_i}(t_m) & \dots & \mu_{c_n}(t_m) \end{pmatrix}. \quad (9)$$

Each row  $\mathbf{R}_j$  in the association matrix  $\mathbf{R}$  is a vector corresponding to the membership degree of a term  $t_j$  belonging to each concept  $c_i$  of the pruned concept space  $C$ ,  $\forall c_i \in C|_{i=1,\dots,n}$ :

$$\mathbf{R}_j^T = (\mu_{c_1}(t_j) \quad \mu_{c_2}(t_j) \quad \dots \quad \mu_{c_i}(t_j) \quad \dots \quad \mu_{c_n}(t_j)). \quad (10)$$

Similarly, each column  $\mathbf{R}_i$  in  $\mathbf{R}$  is a vector corresponding to a concept  $c_i$  giving various degrees of each term  $t_j$ ,  $\forall t_j \in A|_{j=1,\dots,m}$ :

$$\mathbf{R}_i = (\mu_{c_i}(t_1) \quad \mu_{c_i}(t_2) \quad \dots \quad \mu_{c_i}(t_j) \quad \dots \quad \mu_{c_i}(t_m)). \quad (11)$$

By SVD,  $\mathbf{R}$  can be decomposed into the product of three other matrices:

$$\mathbf{R} = \mathbf{U}\mathbf{S}\mathbf{V}^T, \quad (12)$$

where  $\mathbf{S}$  is a  $l \times l$  diagonal matrix such that  $\mathbf{S} = [\delta_{i,j}]$ , where  $\forall i=j \delta_{i,j} \neq 0$  and  $\forall i \neq j \delta_{i,j} = 0$ ,  $\mathbf{U}$  and  $\mathbf{V}$  have orthogonal and unitary columns such that  $\mathbf{U}^T\mathbf{U} = \mathbf{I}$ ,  $\mathbf{V}^T\mathbf{V} = \mathbf{I}$ ,  $\mathbf{I}$  is the identity matrix.

$\mathbf{R}$  can be expressed as

$$\mathbf{R} = ([\mathbf{u}_1] \quad \dots \quad [\mathbf{u}_l]) \cdot \begin{pmatrix} \delta_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \delta_l \end{pmatrix} \cdot \begin{pmatrix} [\mathbf{v}_1] \\ \vdots \\ [\mathbf{v}_l] \end{pmatrix}. \quad (13)$$

$\mathbf{U} = ([\mathbf{u}_j])|_{j=1,\dots,l}$  and  $\mathbf{V} = ([\mathbf{v}_i])|_{i=1,\dots,l}$  are the left and right singular vectors, respectively, corresponding to the monotonically decreasing singular values  $\delta_l|l \in (i, j)$ ,  $i = j$  of the diagonal matrix  $\mathbf{S}$  and  $l = \min(m, n)$ . The full-rank or  $l$ -rank  $\mathbf{R}$  could then be approximated to a rank- $k$  approximation using Latent Semantics Structure [10] in which the largest  $k$  singular values of  $\mathbf{R}$  associated with the first  $k$  columns of the  $\mathbf{U}$  and  $\mathbf{V}$  matrices are used for reconstruction, i.e.,  $\mathbf{R}_k = \mathbf{U}_k \mathbf{S}_k \mathbf{V}_k^T$ , where  $k \ll l$ . Therefore, the  $\mathbf{R}_k$  is the closest rank- $k$  approximation in the least square error sense as it consists of the  $k$  largest singular triplets of  $\mathbf{R}$  [51]. As a result, a new set of membership value  $\mu_{c_i}(t_j)$  will then represent the degree of a term  $t_j$  in the reduced term space with only  $k$ -dimensions, for the concepts  $c_i \in \{c_1, \dots, c_g\}$ . The  $k$  value is empirically determined with respect to the ontology F-measure as defined in Section 6.1.

#### 4.4 Fuzzy Relation Extraction

The final stage toward our ontology extraction method is fuzzy taxonomy generation based on the subsumption relations among the extracted concepts. Let  $Spec(c_x, c_y)$  denote that concept  $c_x$  is a specialization (subclass) of another concept  $c_y$ . The degree of such a specialization relation can be estimated from

$$\begin{aligned} \mu_{RCC}(c_x, c_y) &\approx Spec(c_x, c_y) \\ &= \frac{\sum_{t \in c_x \cap c_y} \mu_{c_x}(t) \otimes \mu_{c_y}(t)}{\sum_{t_x \in c_x} \mu_{c_x}(t_x)}, \end{aligned} \quad (14)$$

where  $\otimes$  is a fuzzy conjunction operator which is equivalent to the min function. The above formula states that the degree of subsumption (specificity) of  $c_x$  to  $c_y$  is based on the ratio of the sum of the minimal membership values of the common terms belonging to both concepts to the sum of the membership values of terms in the concept  $c_x$ . For instance, if every attribute of  $c_y$  is also an attribute of  $c_x$ , a strong specificity relation exists and the value of  $Spec(c_x, c_y)$  is high. The range of the  $Spec(c_x, c_y)$  relation falls in the unit interval  $[0, 1]$  and the subsumption relation is asymmetric. Equation (14) has been applied to our earlier studies of fuzzy ontology extraction [23], [24], [22].

One problem of the standard fuzzy conjunction operation is that the specificity value is highly influenced by the weakest terms (attributes) of the concepts. Therefore, we explore another alternative of estimating the degree of subsumption between two concepts based on the method successfully applied to image analysis [50]. In particular, any two concepts  $c_x$  and  $c_y$  could be said to be similar if their structural similarity is high and the corresponding structural similarity value  $SSIM(c_x, c_y)$  approaches 1 [50]. On the other hand, two concepts are dissimilar if their structural similarity value  $SSIM(c_x, c_y)$  is low (e.g., close to zero). The  $SSIM(c_x, c_y)$  function is expressed as

$$SSIM(c_x, c_y) = l(c_x, c_y) \cdot c(c_x, c_y) \cdot s(c_x, c_y), \quad (15)$$

$$l(c_x, c_y) = \frac{2M_{c_x}M_{c_y} + Q_1}{M_{c_x}^2 + M_{c_y}^2 + Q_1}, \quad (16)$$

$$c(c_x, c_y) = \frac{2\sigma_{c_x}\sigma_{c_y} + Q_2}{\sigma_{c_x}^2 + \sigma_{c_y}^2 + Q_2}, \quad (17)$$

$$s(c_x, c_y) = \frac{\sigma_{c_x c_y} + Q_3}{\sigma_{c_x}\sigma_{c_y} + Q_3}, \quad (18)$$

where  $c_x, c_y \in C$ . For our application, the  $l(c_x, c_y)$  function is used to measure the similarity of two concepts in terms of *semantic coherence*, whereas the  $c(c_x, c_y)$  function is used to estimate the similarity between two concepts in terms of *semantic variance*. Finally, the  $s(c_x, c_y)$  function is applied to measure the similarity of two concepts based on their component structures. Slightly different from [50], our similarity metric is applied to the concept vectors

$c_i = \langle \mu_{c_i}(t_1), \dots, \mu_{c_i}(t_m) \rangle$ . The mean and standard deviation of each concept vector, and the covariance between two concept vectors are defined by

$$M_{c_i} = \frac{1}{m} \left( \sum_{j=1}^m \mu_{c_i}(t_j) \right), \quad (19)$$

$$\sigma_{c_i} = \left( \frac{1}{m-1} \left( \sum_{j=1}^m (\mu_{c_i}(t_j) - M_{c_i})^2 \right) \right)^{1/2}, \quad (20)$$

$$\sigma_{c_x, c_y} = \left( \frac{1}{m-1} \left( \sum_{j=1}^m (\mu_{c_x}(t_j) - M_{c_x})(\mu_{c_y}(t_j) - M_{c_y}) \right) \right). \quad (21)$$

The terms  $Q_1 = 0.0255$ ,  $Q_2 = 0.2295$ , and  $Q_3 = 0.1148$  are constants, and they are applied to image analysis work before [50]. We adopt  $Q_1 = [0, 5 \times 0.0255]$ ,  $Q_2 = [0, 5 \times 0.2295]$ ,  $Q_3 = [0, 5 \times 0.1148]$  in our experiments. When we apply the structural similarity measure to estimate the degree of subsumption between two concepts, we follow the same intuition illustrated in Definition 4. For instance, if most attributes  $t_i$  belonging to the concept  $c_y$  are also belonging to the concept  $c_x$ , the concept  $c_x$  is a subconcept of  $c_y$  to a high degree. To formulate our  $Spec(c_x, c_y)$  function based on the structural similarity, we first compute the common concept  $c_g = c_x \cap c_y$ . Then, we examine if this common subconcept is more subsumed by which concept to determine the direction of the specialization relation. Thereby, the degree of specificity from  $c_x$  to  $c_y$  is approximated by

$$\begin{aligned} \mu_{R_{CC}}(c_x, c_y) &\approx Spec(c_x, c_y) \\ &= \begin{cases} 0 & \text{if } SSIM(c_y, c_g) > SSIM(c_x, c_g), \\ \frac{SSIM(c_x, c_g) - SSIM(c_y, c_g)}{SSIM(c_x, c_g)} & \text{otherwise.} \end{cases} \end{aligned} \quad (22)$$

The above formula states that the degree of subsumption (specificity) of  $c_x$  to  $c_y$  is based on the ratio of the difference between the structural similarity of  $SSIM(c_x, c_g)$  and  $SSIM(c_y, c_g)$  to the normalization factor  $SSIM(c_x, c_g)$ . On the other hand, if more common structural elements are found in  $c_y$  rather than  $c_x$  (i.e.,  $c_y$  is a subconcept of  $c_x$ ), the degree of the specificity relation from  $c_x$  to  $c_y$  is zero.

#### 4.5 Fuzzy Taxonomy Extraction

When a fuzzy taxonomy is built, we only select the subsumption relations such that  $Spec(c_x, c_y) \geq Spec(c_y, c_x)$  and  $Spec(c_x, c_y) > \lambda$ , where  $\lambda$  is a threshold to distinguish significant subsumption relations. The parameter  $\lambda$  is estimated based on empirical tests. If  $Spec(c_x, c_y) \simeq Spec(c_y, c_x)$  and  $Spec(c_x, c_y) > \lambda$  is established, the *equivalent* relation between  $c_x$  and  $c_y$  will be extracted. In addition, a pruning step is introduced such that the redundant taxonomy relations are removed. If the membership of a relation  $\mu_{C \times C}(c_1, c_2) \leq \min(\{\mu_{C \times C}(c_1, c_i), \dots, \mu_{C \times C}(c_i, c_2)\})$ , where  $c_1, c_i, \dots, c_2$  form a path  $P$  from  $c_1$  to  $c_2$ , the relation  $R(c_1, c_2)$  is removed because it can be derived from other stronger specificity relations in the ontology. The fuzzy domain ontology mining algorithm is summarized and shown in Fig. 4. According to this algorithm, more than one

#### Algorithm FuzzyOntoExt( $D, PA, Ont$ )

**Input:** corpus  $D$  and vector of threshold values  $PA$

**Output:** a light weight fuzzy domain ontology  $Ont$

**Main Procedure:**

- 1)  $Ont = \langle \rangle$
- 2) For each document  $d \in D$  Do
  - a) Construct text windows  $w \in d$
  - b) Remove stop words  $sw$  from  $w$
  - c) Perform POS tagging for each term  $t_i \in w$
  - d) Apply Porter stemming to each term  $t_i$
  - e) Filter specific linguistic patterns such as NN, AN, NVN, etc.
  - f) Accumulate the frequency for  $t_i \in w$  and the joint frequency for any pair  $t_i, t_j \in w$
  - g) IF  $lower \leq Freq(t_i) \leq upper$ , THEN  $A = A \cup t_i$
- 3) For each term  $t_i \in A$  Do /\* Concept Extraction \*/
  - a) compute its context vector  $c_i$  using BMI, MI, JA, CP, KL, ECH, or NGD
  - b)  $C = C \cup c_i$
- 4) For each  $c_i \in C$  Do /\* Concept Filtering -  $\alpha$ -cut \*/
  - a) IF  $\exists t_i \in c_i : \mu_{c_i}(t_i) < \zeta$ , THEN  $C = C - c_i$
- 5) For each  $c_i \in C, t_i \in A$  Do /\* Update  $R_{AC}$  relations \*/
  - a) IF  $\mu_{c_i}(t_i) \geq \zeta$ , THEN  $R_{AC} = R_{AC} \cup (t_i, c_i)$
- 6)  $\forall c_i \in C$  : Compute  $Rel(c_i, D_j)$
- 7) IF  $Rel(c_i, D_j) < \varpi$ , THEN  $C = C - c_i$  /\* Concept Pruning \*/
- 8) Perform Dimensionality Reduction SVD
- 9) For each pair of concepts  $(c_i, c_j) \in C$  Do
  - a) Compute the taxonomy relation  $(c_i, c_j)$  using  $Spec(c_i, c_j)$
  - b) IF  $\mu_{R_{CC}}(c_i, c_j) > \lambda$ , THEN  $R_{CC} = R_{CC} \cup (c_i, c_j)$
- 10) For each  $(c_i, c_j) \in R_{CC}$  Do /\* Taxonomy Pruning \*/
  - a) IF  $\mu_{R_{CC}}(c_i, c_j) < \mu_{R_{CC}}(c_j, c_i)$ , THEN  $R_{CC} = R_{CC} - (c_i, c_j)$
  - b) IF  $\exists P(c_i \rightarrow c_x, \dots, c_y \rightarrow c_j)$  AND  $\mu_{R_{CC}}(c_i, c_j) \leq \min(\{\mu_{R_{CC}}(c_i, c_x), \mu_{R_{CC}}(c_x, c_y), \dots, \mu_{R_{CC}}(c_y, c_j)\})$ , THEN  $R_{CC} = R_{CC} - (c_i, c_j)$
- 11)  $Ont = \langle A, C, R_{AC}, R_{CC} \rangle$
- 12) RETURN  $Ont$

Fig. 4. The fuzzy domain ontology extraction algorithm.

directed graph could be generated from a corpus. Each graph will be used as the basis to generate a concept map. The general computational complexity of our algorithm is characterized by  $O(k^2n + kn^2)$ , where  $k$  is the reduced dimensionality of the term space, and  $n$  is the cardinality of the pruned concepts  $C$ . According to our empirical testings, the setting of  $k \simeq n$  can in general lead to satisfactory performance. Therefore, the actual computational complexity of our algorithm is characterized by  $O(k^3)$ . By controlling the concept pruning threshold  $\varpi$ , we can reduce the concept space. Moreover, we can make a trade-off between computational time and accuracy by tuning  $k$  during dimensionality reduction. As a result, our algorithm can scale up for processing a large number of messages online. We have conducted a field test to demonstrate that our system can run efficiently in practice.



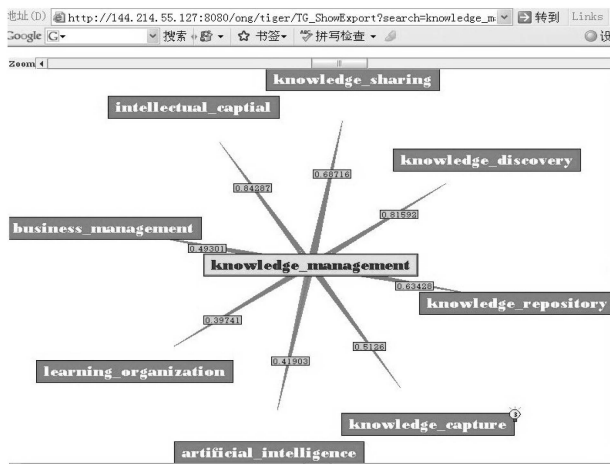


Fig. 5. An automatically generated concept map about “knowledge management.”

## 5 APPLICATION TO E-LEARNING

In an e-Learning environment, learners are often encouraged to reflect what they have learned by writing online journals or sharing their ideas via an online discussion board. Usually, an instructor or other fellow students will reply a new message, and hence, generating multiple threads of messages for each new message posted. If the instructor wants to know the current learning status of their students, she needs to spend a lot of time to browse through all the threads of messages to analyze the contents. Given the fact that humans’ cognitive power is quite limited, such a mental analysis process is very time consuming, and it is very unlikely that the instructor can complete the analysis task on the fly (i.e., when a lecture or tutorial is in progress). To alleviate such a problem and to facilitate adaptive teaching and learning, we can apply the fuzzy domain ontology discovery algorithm to automatically extract and visualize the concept maps representing an individual or a group of learners’ knowledge structure. Based on the concept maps, the instructor can examine whether the existing concepts have been thoroughly internalized by her students or not, and then she can decide which topics should be covered next.

Fig. 5 depicts the concept map about “Knowledge Management,” and the other subconcepts such as knowledge discovery, knowledge capture, intellectual capital, business management, and so on. For readability reason, stemming is not performed for our demonstration examples. The corresponding OWL statements generated by our system are shown in Fig. 6. It should be noted that we use the (rel) attribute of the `<rdfs:comment>` tag to describe the membership of a fuzzy relation (e.g., the superclass/subclass relationship). As the size of each node on the concept map is fixed, some of the characters of the concept labels are truncated (Fig. 5). The number attached to a link connecting each pair of concepts shows the strength of the corresponding subconcept/superconcept relationship. When a node at the second level is clicked, all the subconcepts below the current node will be shown. For instance, when the instructor clicks the “knowledge capture” node (i.e., the node with the number “3” on the top right-hand corner), the subconcepts under this node will be displayed as demonstrated in Fig. 7. The number

```
<?xml version="1.0" ?>
- <rdf:RDF xmlns:rdf="http://ebiz.is.cityu.edu.hk/raylau/ong/rdf" xmlns:xsd=
- <owl:Ontology rdf:about="">
- <rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string":
- <rdfs:label xml:lang="en" />
- <owl:Ontology>
- <owl:Class rdf:ID="business_management">
- <rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string":
- ||rel = 0.49301 </rdfs:comment>
- <rdfs:subClassOf>
- <owl:Class rdf:ID="knowledge_management" />
- <rdfs:subClassOf>
- <owl:Class>
- <owl:Class rdf:ID="knowledge_repository">
- <rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string":
- ||rel = 0.63428 </rdfs:comment>
- <rdfs:subClassOf rdf:resource="#knowledge_management" />
- <owl:Class>
- <owl:Class rdf:ID="learning_organization">
- <rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string":
- <rdfs:subClassOf rdf:resource="#knowledge_management" />
```

Fig. 6. A segment of OWL statements for the “knowledge management” ontology.

attached to the top right-hand corner of a node indicates the number of levels below the current node. The user is provided with the option to save the concept map and the corresponding OWL file to a local disk to facilitate learning from the peers.

As a fuzzy domain ontology may contain hundreds of nodes, it may be difficult to display all the concepts in one single concept map. We adopt a linearization procedure [53] to generate the concepts map for each main concept when the number of nodes in an ontology exceeds 100. Basically, each node can be seen as the root concept of an individual concept map. We are examining using other tools such as OntoSphere [2] to visualize a complex ontology using a 3D display. Our prototype system was developed using Java (J2SE v 1.4.2), Java Server Pages (JSP) 2.1, and Servlet 2.5. For the implementation of SVD for our term space reduction, we employed the publicly available Java toolkit called GAP.<sup>3</sup> For the visualization of the concept maps, we developed our visualization module based on the Java-based shareware TouchGraph.<sup>4</sup> Our prototype system is operated under Apache Tomcat 6.0 Web server.<sup>5</sup>

## 6 SYSTEM EVALUATION

### 6.1 Evaluation Metrics

We try to evaluate the automatically generated concept maps by comparing them with the maps developed by human experts. Our first evaluation metric is developed based on the Generalized Distance Ratio (GDR) method [29]. The GDR measure is the generalization of Langfield-Smith and Wirth’s metric [20], and it has been widely used to quantitatively evaluate concept maps in the fields of education, operational research, and strategic management [29]. The GDR measure aims at comparing concept maps by using all the available information encoded in the maps. Specifically, the GDR measure considers three types of difference: 1) existence or nonexistence of elements (nodes), 2) existence and nonexistence of beliefs (arcs), and

3. <http://gap.stat.sinica.edu.tw/Software/GAP/>.

4. <http://sourceforge.net/projects/touchgraph>.

5. <http://144.214.55.127:8080/ong/tiger/OntologyUI.jsp>.

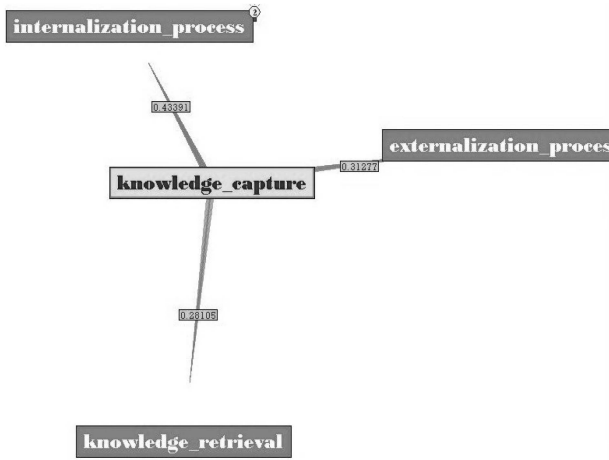


Fig. 7. The display of subconcept “knowledge capture.”

3) identical beliefs (arcs) with different strengths (i.e., the membership of our fuzzy relation  $R_{CC}$ ).

Originally, the GDR metric has five parameters such as  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\varepsilon$ , and  $\delta$  to deal with different kinds of map comparisons [29]. We contextualize the GDR metric to meet our specific map comparison requirements by employing the following parameter values:  $\alpha = 1$  (no account for node self-influence),  $\beta = \varepsilon = 1$  (the weights of arcs falling in the unit interval  $[0, 1]$ ),  $\gamma = 1$  (two arcs are considered different if the connecting pair of nodes are not the same),  $\delta = 0$  (no consideration of polarity change). The adapted metric values fall in the unit interval for the comparison of any two maps. The adapted Generalized Distance Ratio (DR) measure is defined as follows:

$$DR(A, B) = \frac{\sum_{i=1}^p \sum_{j=1}^p diff(i, j)}{p_c^2 + 2p_c(p_{uA} + p_{uB}) + p_{uA}^2 + p_{uB}^2 - (p_c + p_{uA} + p_{uB})}, \quad (23)$$

$$diff(i, j) = \begin{cases} 0 & \text{if } i = j, \\ 1 & \text{if } (a_{ij} \neq 0 \vee b_{ij} \neq 0) \\ & \wedge (i \notin P_c \vee j \notin P_c), \\ |a_{ij} - b_{ij}| & \text{otherwise,} \end{cases} \quad (24)$$

where  $A$  and  $B$  are two extended adjacency matrices of size  $p$ . The term  $a_{ij}$  (or  $b_{ij}$ ) is the value of the  $i$ th row and  $j$ th column of ( $A$  or  $B$ ),  $P_c$  is the set of nodes common to both maps,  $p_c = |P_c|$  is the cardinality of the set  $P_c$ ,  $p_{uA}$  is the number of unique nodes of map  $A$ , and  $p_{uB}$  is the number of unique nodes of map  $B$ .

We also employ standard measures such as recall, precision, and the F-measure developed in the field of IR [39] to evaluate the concept maps. In particular, we develop ontology recall  $Ont\_Recall$ , ontology precision  $Ont\_precision$ , and ontology F-measure  $Ont\_F$  as follows:

$$Node\_Recall = \frac{|N_{ME} \cap N_{MS}|}{|N_{ME}|}, \quad (25)$$

$$Node\_Precision = \frac{|N_{ME} \cap N_{MS}|}{|N_{MS}|}, \quad (26)$$

$$Link\_Recall = \frac{|L_{ME} \cap L_{MS}|}{|L_{ME}|}, \quad (27)$$

$$Link\_Precision = \frac{|L_{ME} \cap L_{MS}|}{|L_{MS}|}, \quad (28)$$

$$Ont\_Recall = \omega_R \times Node\_Recall + (1 - \omega_R) \times Link\_Recall, \quad (29)$$

$$Ont\_Precision = \omega_P \times Node\_Precision + (1 - \omega_P) \times Link\_Precision, \quad (30)$$

$$F_\eta = \frac{(1 + \eta^2) Precision \times Recall}{\eta^2 Precision + Recall}, \quad (31)$$

where  $N_{ME}$  and  $N_{MS}$  represent the set of nodes found from the concept map created by human experts and that generated by our system, respectively. Similarly,  $L_{ME}$  and  $L_{MS}$  are the set of links encoded on the concept map drawn by human experts and the concept map generated by our system, respectively. In fact, the set of links can easily be identified from the adjacency matrices which encode concept maps. In particular, only the upper half or the lower half of each matrix needs to be traversed to construct a link set. The parameter  $\omega_R$  is used to compute the ontology recall based on a weighted sum of the node recall and link recall, respectively. Similarly,  $\omega_P$  is used to tune the ontology precision measure. For the experiments presented in this paper, we adopt  $\omega_R = \omega_P = 0.5$ . The standard F-measure is shown in (31) [49]. If we assume that precision is as important as recall (i.e.,  $\eta = 1$ ), the ontology F-measure  $Ont\_F$  is defined by

$$Ont\_F = \frac{2 \times Ont\_Precision \times Ont\_Recall}{Ont\_Precision + Ont\_Recall}. \quad (32)$$

## 6.2 Benchmark Tests

The TREC-AP, which comprises the Associated Press (AP) newswires covering the period from 1988 to 1990, is the benchmark corpus used for our experiments [37]. We used a TREC topic description and at most five relevant TREC-AP documents associated with the particular topic to simulate the online messages generated from an e-Learning platform. A human expert (a postdoctoral researcher in the field of banking and finance) was recruited to read the topic description as well as the associated relevant documents so that she could draw a concept map illustrating the main concepts and propositions for each of the TREC topic. These concept maps become the benchmark for comparison with the system generated concept maps based on the metrics developed in Section 6.1. For each experimental run, we manipulated different system parameters to test different aspects of our system. We selected TREC topics 1 to 10 and topics 41 to 50 for our experiments since each of these topics has at least five relevant documents. For instance, TREC topic 49 is about “who’s working with Supercomputers.”

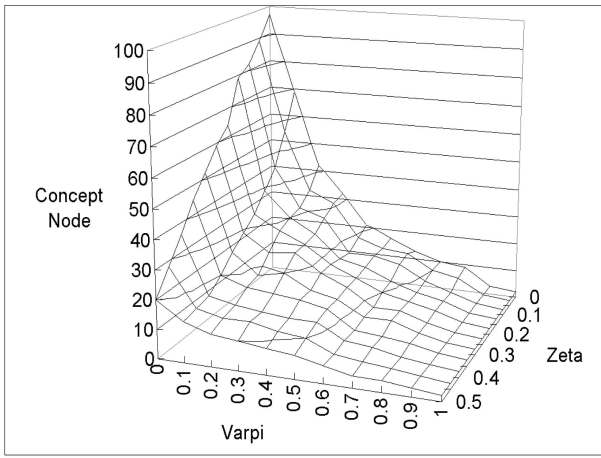


Fig. 8. Average number of concepts generated by controlling  $\zeta$  and  $\varpi$ .

**Experiment 1.** The purpose of the first experiment is to test the effectiveness of the concept extraction/pruning thresholds. We used the BMI method (2) for concept extraction and the standard fuzzy conjunction operation (14) for fuzzy relation extraction. Other parameters included  $\beta = 0.672$  [22] and  $\lambda = 0.093$ . The noun phrase patterns “Noun Noun” were used for all the experiments discussed in this paper. We used other five domains (entertainment, education, humanity, sport, and arts) as the basis to compute the concept relevance scores during concept pruning. Each domain consists of the first 1,000 Web pages retrieved via the Google Search API.<sup>6</sup> When the domain frequency  $Dom(c_i, D_k)$  was calculated for the TREC-AP domain, we converted the document basis to 1,000 as well. The average number of concept nodes generated and the average ontology F-measure achieved over the 20 TREC domains under various combinations of  $\zeta$  and  $\varpi$  are plotted in Figs. 8 and 9, respectively. As shown in Fig. 8, when both  $\zeta \geq 0.2$  and  $\varpi \geq 0.4$  were applied, the number of concept nodes generated by our system would be reduced dramatically. It indicates that our concept pruning mechanism can work effectively. From Fig. 9, it is shown that the best ontology F-measure could be achieved if  $\zeta$  and  $\varpi$  are set to the ranges  $[0.35, 0.45]$  and  $[0.4, 0.6]$ , respectively. The reason is that most noisy concepts will be filtered out under such a combination; setting higher threshold values could not lead to the best performance either because many important domain concepts were pruned.

**Experiment 2.** The objective of this experiment is to evaluate the effectiveness of different concept extraction methods such as BMI, JA, CP, KL, ECH, and NGD illustrated in Section 4.1. When different methods are applied, the underlying terms and the term weights  $\mu_{c_i}(t)$  associated with a concept may be different. Such a difference can be realized when we apply (22) or (14) to compute the fuzzy relations between concepts because both of the metrics will compare the concepts based on their underlying semantics (e.g., the composing terms and their weights). We adopted the same system parameters as in experiment one. The average link precisions achieved by various concept extraction methods under different extraction threshold values  $\zeta = [0, 0.5]$  are

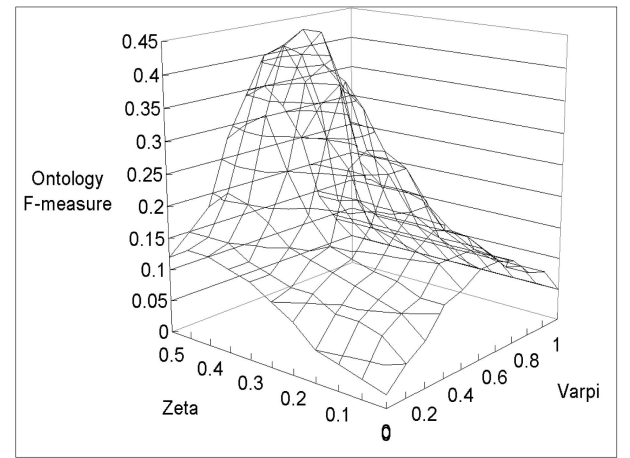


Fig. 9. The average F-measure by tuning  $\zeta$  and  $\varpi$ .

plotted in Fig. 10. In general, the link precision is improved when higher concept extraction threshold is used because less noisy terms will be used to construct the corresponding concept vectors. As shown in Fig. 10, the BMI method outperforms the other methods in terms of average link precision most of the time (e.g., from  $\zeta = 0.25$  to  $\zeta = 0.5$ ).

**Experiment 3.** We also examined the effectiveness of (22) and (14) which were used to estimate the strength of a concept specialization relation  $\mu_{RCC}(c_x, c_y)$  given any two concepts  $c_x, c_y$ . In this experiment, we used the BMI concept extraction method and we set the parameters  $\zeta = 0.431$  and  $\varpi = 0.512$ ; other parameters remained the same as before. In the first run, we used the standard fuzzy conjunction operator (14) for concept map generation; the second run involved the structural similarity SSIM method (22) under the same conditions. The parameters of the SSIM method were  $Q_1 = 0.026$ ,  $Q_2 = 0.459$ , and  $Q_3 = 0.344$ . A topic-by-topic comparison in terms of ontology precision, ontology recall, ontology F-measure, and DR are tabulated in Tables 1 and 2, respectively. The second and the third columns refer to the number of concept nodes and concept relations generated by the system. By testing the hypotheses:  $H_{Null} : \mu_{SSIM} - \mu_{Fuzzy} = 0$  and  $H_{Alternative} : \mu_{SSIM} - \mu_{Fuzzy} > 0$  with

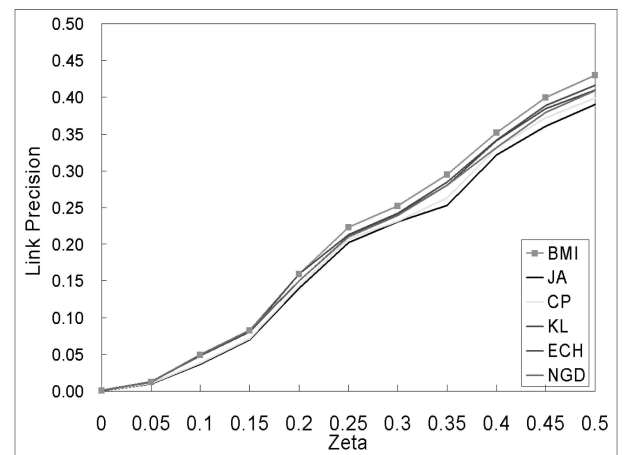


Fig. 10. The relative link precision of various concept extraction methods.

6. <http://code.google.com/apis/soapsearch/>.

TABLE 1  
Concept Map Generation Using Standard Fuzzy  
Conjunction Operator

Topic	Node	Link	Ontology Recall	Ontology Precision	Ontology F-measure	DR
1	25	116	0.763	0.474	0.585	0.267
2	27	95	0.733	0.568	0.640	0.215
3	22	91	0.691	0.615	0.651	0.199
4	18	90	0.691	0.522	0.595	0.244
5	16	88	0.765	0.409	0.533	0.324
6	24	115	0.673	0.287	0.402	0.418
7	27	131	0.688	0.419	0.521	0.297
8	18	93	0.754	0.527	0.620	0.213
9	19	87	0.712	0.540	0.614	0.254
10	22	98	0.731	0.388	0.507	0.387
41	16	89	0.634	0.292	0.400	0.318
42	24	81	0.625	0.370	0.465	0.329
43	23	85	0.711	0.694	0.702	0.157
44	19	96	0.600	0.375	0.462	0.330
45	20	97	0.727	0.412	0.526	0.279
46	22	109	0.754	0.395	0.518	0.303
47	18	89	0.736	0.315	0.441	0.321
48	19	85	0.667	0.353	0.462	0.362
49	16	101	0.683	0.277	0.394	0.346
50	15	87	0.806	0.333	0.472	0.356
Avg.	20.5	96.15	0.707	0.428	0.526	0.296

paired one-tail  $t$ -test on the F-measure scores obtained from the 20 TREC-AP topics, the null hypothesis is rejected ( $t(19) = 3.067, p < 0.01$ ). Therefore, it is confirmed that the SSIM method (22) for concept relation extraction is more effective than the method using standard fuzzy conjunction operator (14). As can be seen, the average distance between the maps generated by our system and the maps drawn by the domain expert is 0.285 only. It means that the concept maps produced by our system closely resemble the maps constructed by the domain expert.

### 6.3 Field Tests

Field tests were conducted to verify the quality of the concept maps generated by our system. The subjects were a group of postgraduate students taking a Knowledge Management course. These subjects learned about concept mapping in their classes. At the end of a lecture, subjects were told to reflect the main concepts they learned from the class by writing short messages on an online discussion forum. The time given to them to write the messages was limited to 10 minutes for each class. After the subjects had finished their reflection, the concept map generation tool was invoked to automatically construct the concept maps representing the group's perception about the concepts covered in the lecture. We employed the BMI method for concept extraction and the SSIM method for relation extraction. Other system parameters were the same as those used in experiment three. Each subject was given another 10 minutes to browse through the concept maps generated by the system, and then they would answer a questionnaire. Our questionnaire was developed based on the instrument employed in [5]. It included the assessment of five factors: Accuracy, Cohesiveness, Isolation, Hierarchy, and Readability.

A five-point semantic differential scale from very good (5), good (4), average (3), bad (2), to very poor (1) is used to measure the dependent variables. In general, a score close to 5 indicates that the automatically generated concept

TABLE 2  
Concept Map Generation Using SSIM

Topic	Node	Link	Ontology Recall	Ontology Precision	Ontology F-measure	DR
1	25	117	0.748	0.464	0.572	0.272
2	27	97	0.721	0.536	0.615	0.231
3	22	90	0.704	0.633	0.667	0.190
4	18	93	0.676	0.495	0.571	0.258
5	16	85	0.766	0.424	0.545	0.316
6	24	116	0.633	0.267	0.376	0.429
7	27	125	0.738	0.472	0.576	0.270
8	18	90	0.769	0.556	0.645	0.199
9	19	88	0.727	0.545	0.623	0.251
10	22	93	0.750	0.419	0.538	0.367
41	16	86	0.683	0.329	0.444	0.302
42	24	74	0.687	0.446	0.541	0.289
43	23	82	0.735	0.744	0.739	0.131
44	19	99	0.583	0.354	0.440	0.342
45	20	86	0.764	0.488	0.596	0.243
46	22	101	0.772	0.436	0.557	0.283
47	18	86	0.763	0.337	0.468	0.310
48	19	83	0.689	0.373	0.484	0.351
49	16	93	0.732	0.323	0.448	0.324
50	15	82	0.806	0.354	0.492	0.345
Avg.	20.5	93.25	0.722	0.449	0.547	0.285

map is with good quality and it can reflect what the group perceived about the subject topic. The results of the field tests are shown in Table 3. The second column indicates the number of subjects involved in a field test, the third and the fourth columns show the number of concepts nodes and links automatically generated by the system, and the fifth column shows the time (in minutes) spent on generating the concept maps. The overall mean scores for accuracy, cohesiveness, isolation, hierarchy, and readability are 4.23, 4.22, 4.15, 4.31, and 3.95, respectively. For most of the dependent variables, the overall mean score is above 4 except the readability issue. The reason for a bit lower score in readability may be that our programmer used a small fixed size rectangle to represent concept node. The time taken to generate the concept map (including the underlying OWL statements) on our Web server varied from 1.3 to 1.8 minutes. This result indicates that it is feasible for instructors to invoke such a tool to analyze students' understanding about a subject topic on the fly.

## 7 RELATED RESEARCH

There is a large number of educational intermediaries storing metadata descriptions for various learning resources to facilitate educational knowledge management [34]. In order to ensure effective communications between the users and the learning resources, automatic discovery of the taxonomies of these learning resources is required. A data mining approach was proposed to discover the relations of the metadata describing the various learning resources. A graph-based clustering algorithm was applied to extract meaningful concepts for the learning resources and to identify the relations among the concepts. Our work aims at extracting and visualizing the concept maps based on the online messages created by the learners rather than discovering the ontology of educational resources. We employ a hybrid lexico-syntactic and statistical learning method rather than a computationally expensive graph-based approach for ontology extraction.

TABLE 3  
The Results of Field Tests

Lecture	Subject	Node	Link	Time	Accuracy		Cohesiveness		Isolation		Hierarchy		Readability	
					Mean	STD	Mean	STD	Mean	STD	Mean	STD	Mean	STD
Knowledge Management	20	16	61	1.5	4.20	0.871	4.25	0.698	4.05	0.865	4.41	0.663	3.93	0.831
Knowledge Discovery	19	17	63	1.6	4.21	0.885	4.22	0.673	4.15	0.763	4.32	0.822	4.01	0.642
Knowledge Sharing	22	21	79	1.8	4.25	0.626	4.17	0.759	4.16	0.833	4.28	0.682	3.87	0.693
Knowledge Capture	20	18	55	1.5	4.33	0.678	4.35	0.622	4.29	0.654	4.29	0.731	4.13	0.834
Knowledge Application	18	16	59	1.3	4.17	0.654	4.09	0.827	4.11	0.846	4.23	0.693	3.82	0.675
Average		17.60	63.40	1.54	4.23	0.743	4.22	0.716	4.15	0.792	4.31	0.718	3.95	0.735

There was also research work discussing the general ideas of automatically extracting ontologies from teaching documents although the actual implementation of the proposal was missing [18]. Previous work had also employed the Term Frequency Inverse Document Frequency (TFIDF) heuristic developed from the field of IR to extract prominent concepts from electronic messages generated in e-Learning [52]. A knowledge density score was developed based on the TFIDF term weighting formula to assess the extent of contribution to online knowledge sharing by individuals. Our document parsing approach also employs TFIDF and other linguistic pattern recognition method to extract concepts from text. In addition, we deal with the automatic construction of a taxonomy of concepts as well.

The FOGA framework for fuzzy ontology extraction has been reported [48]. The FOGA framework consists of fuzzy formal concept analysis, fuzzy conceptual clustering, fuzzy ontology generation, and semantic representation conversion. The notions of formal context and formal concept have been fuzzified by introducing the respective membership functions. The FOGA framework is evaluated in a small citation database. Our method discussed in this paper differs from the FOGA framework in that a more compact representation of fuzzy domain ontology is developed. Our proposed method is based on the theory developed in computational linguistic and our computational algorithm is developed with respect to the concept of fuzzy relations. Our proposed method is validated with respect to a larger benchmark corpus and a field test.

A fuzzy ontology which is an extension of the crisp domain ontology was utilized for news summarization purpose [25]. In this semiautomatic ontology discovery approach, the domain ontology about various events covered by some net news was manually developed by human domain experts. The main function of the automatic fuzzy inference mechanism was to generate the membership degrees (classification) for each event with respect to the fuzzy concepts defined in the fuzzy ontology. The standard triangular membership function was used for the classification purpose. The method discussed in this paper is a fully automatic fuzzy domain ontology discovery approach. There is no predefined fuzzy concepts and taxonomy of concepts, instead our fuzzy domain ontology extraction method will automatically discover the concepts and generate the taxonomy relations.

An ontology mining technique was proposed to extract patterns representing users' information needs [26]. The ontology mining method comprised two parts: the top backbone and the base backbone. The former represented the relations between compound classes of the ontology.

The latter indicated the linkage between primitive classes and compound classes. The Dempster-Shafer theory of evidence model was applied to extract the relations among classes. The research work presented in this paper focuses on fuzzy domain ontology discovery rather than the discovery of crisp ontology representing users' information needs. Instead of using Dempster-Shafer theory of evidence, our concept extraction method is underpinned by information theoretic approaches such as mutual information.

Sanderson and Croft [40] proposed a document-based subsumption induction method to automatically derive a hierarchy of terms from a corpus. In particular, the subsumption relations among terms were developed based on the co-occurrence of terms in the documents of a corpus. An artificial threshold was required to define a cutoff point to determine the specificity relation between terms. Our method differs from their work in that we are dealing with the more challenging task of concept hierarchy extraction rather than term relationship extraction. In addition, our method extends their concept extraction approach in that the co-occurrence of terms is derived based on a moving text window rather than the whole document to reduce the chance of generating noisy subsumption relations.

An ontology discovery approach was proposed to improve domain ontologies by mining the hidden semantics from text [8]. The learning approach was based on self-organizing map (SOM). The SOM approach was illustrated with reference to the tourism domain and a field test based on the largest Austrian tourism Web site was conducted to validate the ideas. Our ontology extraction method is based on context-sensitive text mining and fuzzy relation construction rather than using SOM. Moreover, we employ the notion of fuzzy ontology rather than crisp ontology to explicitly model the uncertainty arising in automated ontology extraction.

## 8 CONCLUSIONS AND FUTURE WORK

With the increasing number of online messages generated from interactive e-Learning environments, instructors are often overwhelmed. As a result, adaptive teaching and learning is difficult to achieve. This paper illustrates a novel concept map generation technique which is underpinned by a context-sensitive text mining method and a fuzzy domain ontology extraction algorithm. The proposed mechanism can automatically construct concept maps based on the messages posted to an online discussion board. With such an intelligent tool, instructors can quickly identify the learning status of their students, and hence, more suitable pedagogy can be developed for the subsequent lessons. Our initial experimental results show that the accuracy and the quality of the automatically generated concept maps are

promising. Future work involves a larger scale of field test for the concept map generation mechanism. Other text mining methods will also be explored to improve our fuzzy domain ontology extraction method.

## ACKNOWLEDGMENTS

The work reported in this paper has been funded in part by the Australian Research Council (ARC) Discovery Grant (DP0556455). This research work was also supported in part by the e-Learning research Grant (6980081-680) of City University of Hong Kong, and the United Kingdom's Engineering and Physical Sciences Research Council (EPSRC) Grant (EP/E002145/1).

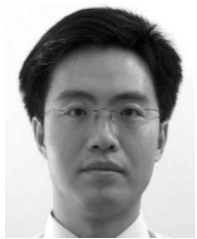
## REFERENCES

- [1] D.M. Bikel, R. Schwartz, and R.M. Weischedel, "An Algorithm that Learns What's in a Name," *Machine Learning*, vol. 34, pp. 211-231, 1999.
- [2] A. Bosca, D. Bonino, and P. Pellegrino, "Ontosphere: More Than a 3D Ontology Visualization Tool," *Proc. Second Italian Semantic Web Workshop Semantic Web Applications and Perspectives (SWAP '05)*, P. Bouquet and G. Tummarello, eds., CEUR-WS.org, Dec. 2005.
- [3] J. Bucur, "HyWrite: Writing in Hypermedia eLearning Environments," *Proc. 17th Conf. Hypertext and Hypermedia (HYPERTEXT '06)*, pp. 45-48, 2006.
- [4] Y.-L. Chi, "Elicitation Synergy of Extracting Conceptual Tags and Hierarchies in Textual Document," *Expert Systems with Applications*, vol. 32, no. 2, pp. 349-357, 2007.
- [5] S. Chuang and L. Chien, "Taxonomy Generation for Text Segments: A Practical Web-Based Approach," *ACM Trans. Information Systems*, vol. 23, no. 4, pp. 363-396, 2005.
- [6] R. Cilibrasi and P. Vitanyi, "The Google Similarity Distance," *IEEE Trans. Knowledge and Data Eng.*, vol. 19, no. 3, pp. 370-383, Mar. 2007.
- [7] P. Cimiano, A. Hotho, and S. Staab, "Learning Concept Hierarchies from Text Corpora Using Formal Concept Analysis," *J. Artificial Intelligence Research*, vol. 24, pp. 305-339, 2005.
- [8] M. Dittenbach, H. Berger, and D. Merkl, "Improving Domain Ontologies by Mining Semantics from Text," *Proc. First Asia-Pacific Conf. Conceptual Modelling (APCCM '04)*, pp. 91-100, 2004.
- [9] G. Forsythe, M. Malcolm, and C. Moler, *Computer Methods for Mathematical Computations*. Prentice Hall, 1977.
- [10] G. Furnas, S. Deerwester, S.T. Dumais, T.K. Landauer, R. Harshman, L.A. Streeter, and K.E. Lochbaum, "Information Retrieval Using a Singular Value Decomposition Model of Latent Semantic Structure," *Proc. 11th Ann. Int'l ACM SIGIR Conf. Research and Development in Information Retrieval (SIGIR '88)*, Y. Chiaramella, ed., pp. 465-480, 1988.
- [11] I. Ganchev, M. O'Droma, and R. Andreev, "Functionality and SCORM-Compliance Evaluation of eLearning Tools," *Proc. Seventh IEEE Int'l Conf. Advanced Learning Technologies (ICALT '07)*, pp. 467-469, 2007.
- [12] P. Gärdenfors, *Conceptual Spaces: The Geometry of Thought*. The MIT Press, 2000.
- [13] T.R. Gruber, "A Translation Approach to Portable Ontology Specifications," *Knowledge Acquisition*, vol. 5, no. 2, pp. 199-220, 1993.
- [14] Z. Harris, *Mathematical Structures of Language*. Wiley, 1968.
- [15] I. Horrocks, P.F. Patel-Schneider, and F. van Harmelen, "From SHIQ and RDF to OWL: The Making of a Web Ontology Language," *J. Web Semantics*, vol. 1, no. 1, 2004.
- [16] L. Howard, Z. Remenyi, and G. Pap, "Adaptive Blended Learning Environments," *Proc. Ninth Int'l Conf. Eng. Education (ICEE '06)*, pp. 11-16, 2006.
- [17] H. Jing and E. Tzoukermann, "Information Retrieval Based on Context Distance and Morphology," *Proc. 22nd Ann. Int'l ACM SIGIR Conf. Research and Development in Information Retrieval (SIGIR '99)*, pp. 90-96, 1999.
- [18] J. Kay and S. Holden, "Automatic Extraction of Ontologies from Teaching Document Metadata," *Proc. Int'l Conf. Computers in Education (ICEE '02)*, pp. 1555-1556, 2002.
- [19] D. Koller and M. Sahami, "Hierarchically Classifying Documents Using Very Few Words," *Proc. 14th Int'l Conf. Machine Learning (ICML '97)*, D.H. Fisher, ed., pp. 170-178, 1997.
- [20] K. Langfield-Smith and A. Wirth, "Measuring Differences between Cognitive Maps," *J. Operational Research Soc.*, vol. 43, no. 12, pp. 1135-1150, 1992.
- [21] R. Lau, "Context-Sensitive Text Mining and Belief Revision for Intelligent Information Retrieval on the Web," *Web Intelligence and Agent Systems: An Int'l J.*, vol. 1, no. 3/4, pp. 1-22, 2003.
- [22] R. Lau, "Fuzzy Domain Ontology Discovery for Business Knowledge Management," *IEEE Intelligent Informatics Bull.*, vol. 8, no. 1, pp. 29-41, 2007.
- [23] R. Lau, A.Y. Chung, D. Song, and Q. Huang, "Towards Fuzzy Domain Ontology Based Concept Map Generation for e-Learning," *Proc. Sixth Int'l Conf. Web-Based Learning (ICWL '07)*, H.L. et al., eds., pp. 90-101, 2007.
- [24] R. Lau, Y. Li, and Y. Xu, "Mining Fuzzy Domain Ontology from Textual Databases," *Proc. IEEE/WIC/ACM Int'l Conf. Web Intelligence (WI '07)*, Nov. 2007.
- [25] C.-S. Lee, Z.-W. Jian, and L.-K. Huang, "A Fuzzy Ontology and Its Application to News Summarization," *IEEE Trans. Systems, Man, and Cybernetics, Part B*, vol. 35, no. 5, pp. 859-880, 2005.
- [26] Y. Li and N. Zhong, "Mining Ontology for Automatically Acquiring Web User Information Needs," *IEEE Trans. Knowledge and Data Eng.*, vol. 18, no. 4, pp. 554-568, Apr. 2006.
- [27] A. Maedche, V. Pekar, and S. Staab, "Ontology Learning Part One—On Discovering Taxonomic Relations from the Web," *Proc. IEEE/WIC/ACM Int'l Conf. Web Intelligence (WI '03)*, N. Zhong, J. Liu, and Y. Yao, eds., pp. 3-24, 2003.
- [28] A. Maedche and S. Staab, "Ontology Learning," *Handbook on Ontologies*, pp. 173-190, 2004.
- [29] L. Markoczy and J. Goldberg, "A Method for Eliciting and Comparing Causal Maps," *J. Management*, vol. 21, no. 2, pp. 305-309, 1995.
- [30] G.A. Miller, B.R.C. Fellbaum, D. Gross, and K.J. Miller, "Introduction to WordNet: An On-Line Lexical Database," *J. Lexicography*, vol. 3, no. 4, pp. 234-244, 1990.
- [31] C.A. Montgomery, "Concept Extraction," *Am. J. Computational Linguistics*, vol. 8, no. 2, pp. 70-73, 1982.
- [32] R. Navigli, P. Velardi, and A. Gangemi, "Ontology Learning and Its Application to Automated Terminology Translation," *IEEE Intelligent Systems*, vol. 18, no. 1, pp. 22-31, Jan./Feb. 2003.
- [33] J. Novak and D. Gowin, *Learning How to Learn*. Cambridge Univ. Press, 1984.
- [34] C. Papatheodorou, A. Vassiliou, and B. Simon, "Discovery of Ontologies for Learning Resources Using Word-Based Clustering," *Proc. World Conf. Educational Multimedia, Hypermedia, and Telecomm. (ED-MEDIA '02)*, pp. 324-326, 2002.
- [35] P. Perrin and F. Petry, "Extraction and Representation of Contextual Information for Knowledge Discovery in Texts," *Information Sciences*, vol. 151, pp. 125-152, 2003.
- [36] M. Porter, "An Algorithm for Suffix Stripping," *Program*, vol. 14, no. 3, pp. 130-137, 1980.
- [37] S. Robertson and I. Soboroff, "The TREC 2001 Filtering Track Report," *Proc. 10th Text REtrieval Conf. (TREC '01)*, E. Voorhees and D. Harman, eds., pp. 26-37, Nov. 2001.
- [38] G. Salton, "Full Text Information Processing Using the Smart System," *Database Eng. Bull.*, vol. 13, no. 1, pp. 2-9, Mar. 1990.
- [39] G. Salton and M. McGill, *Introduction to Modern Information Retrieval*. McGraw-Hill, 1983.
- [40] M. Sanderson and B. Croft, "Deriving Concept Hierarchies from Text," *Proc. 22nd Ann. Int'l ACM SIGIR Conf. Research and Development in Information Retrieval (SIGIR '99)*, pp. 206-213, ACM, 1999.
- [41] H. Schütze, "Automatic Word Sense Discrimination," *Computational Linguistics*, vol. 24, no. 1, pp. 97-124, 1998.
- [42] S. Sekine, J.J. Carroll, S. Ananiadou, and J. Tsujii, "Automatic Learning for Semantic Collocation," *Proc. Third Conf. Applied Natural Language Processing (ANLP '92)*, pp. 104-110, Mar./Apr. 1992.
- [43] C. Shannon, "A Mathematical Theory of Communication," *Bell System Technology J.*, vol. 27, pp. 379-423, 1948.
- [44] M. Sloman, *The e-Learning Revolution: How Technology Is Driving a New Training Paradigm*. AMACOM, 2002.
- [45] G. Smith, *Computers and Human Language*. Oxford Univ. Press, 1991.

- [46] M.H. Soong, H.C. Chan, B.C. Chua, and K.F. Loh, "Critical Success Factors for On-Line Course Resources," *Computers and Education*, vol. 36, no. 2, pp. 101-120, 2001.
- [47] M.A. Stairmand, "Textual Context Analysis for Information Retrieval," *Proc. 20th Ann. Int'l ACM SIGIR Conf. Research and Development in Information Retrieval (SIGIR '97)*, pp. 140-147, 1997.
- [48] Q.T. Tho, S.C. Hui, A.C.M. Fong, and T.H. Cao, "Automatic Fuzzy Ontology Generation for Semantic Web," *IEEE Trans. Knowledge and Data Eng.*, vol. 18, no. 6, pp. 842-856, June 2006.
- [49] C. van Rijsbergen, *Information Retrieval*. Butterworths, 1979.
- [50] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," *IEEE Trans. Image Processing*, vol. 13, no. 4, pp. 600-612, 2004.
- [51] J.H. Wilkinson and C. Reinsch, *Handbook for Automatic Computation: Vol. II. Linear Algebra*. Springer-Verlag, 1971.
- [52] Y. Wu and X. Chen, "eLearning Assessment through Textual Analysis of Class Discussions," *Proc. Fifth IEEE Int'l Conf. Advanced Learning Technologies (ICALT '05)*, pp. 388-390, 2005.
- [53] J. Yeh, C. Wu, M. Chen, and L. Yu, "Automated Alignment and Extraction of a Bilingual Ontology for Cross-Language Domain-Specific Applications," *Computational Linguistics and Chinese Language Processing*, vol. 10, no. 1, pp. 35-52, 2005.
- [54] L.A. Zadeh, "Fuzzy Sets," *J. Information and Control*, vol. 8, pp. 338-353, 1965.



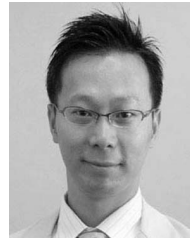
**Raymond Y.K. Lau** received the MAppSc degree in information studies, the MSc degree in systems analysis and design, and the PhD degree in information technology. He is an assistant professor in the Department of Information Systems, City University of Hong Kong, Kowloon, Hong Kong. He has authored more than 60 refereed international journals and conference papers. His research work is published in renowned journals such as the *ACM Transactions on Information Systems*, the *Journal of Management Information Systems*, *Decision Support Systems*, and the *Journal of Pattern Recognition and Artificial Intelligence*. His research interests include information retrieval, text mining, e-Learning, and agent-mediated e-Commerce. He is a senior member of the ACM and the IEEE. He is the associate editor of the *International Journal of Systems and Service-Oriented Engineering*.



**Dawei Song** received the PhD degree from the Chinese University of Hong Kong in 2000. He is a full professor in the School of Computing at Robert Gordon University, United Kingdom. His major research interests include information retrieval models, Web and Enterprise search systems, and biomedical information retrieval. His work has been published in renowned journals such as the *ACM Transactions on Information Systems*, the *Journal of American Society for Information Science and Technology*, and *Decision Support Systems*, and top conferences such as ACM SIGIR and CIKM. He is on the editorial board of the *International Journal of Computer Processing of Oriental Languages*, and a guest editor for the *ACM Transactions on Asian Language Information Processing*.



**Yuefeng Li** received the BSc degree in mathematics and the MSc degree in computer science from Jilin University, China, and the PhD degree in computer science from Deakin University, Australia. He is an associate professor in the Faculty of Information Technology, Queensland University of Technology, Brisbane, Australia. His current research interests include Web mining and Web personalization, ontology learning, Web intelligence, and data mining. He has published more than 100 refereed papers and five books. His research work has appeared and frequently cited in top journals (*IEEE TKDE*, *JMIS*, *DSS*, *KBS*, and *AAI*) and international conferences (ICDM, CIKM, Web Intelligence, and RSCTC). He is an associate editor of the *International Journal of Pattern Recognition and Artificial Intelligence* and an associate editor of the *IEEE Intelligent Informatics Bulletin*. He is a member of the IEEE.



**Terence C.H. Cheung** received the BEng (Hons) degree in computer engineering and the PhD degree in electronic engineering from the City University of Hong Kong, Kowloon, Hong Kong, in 1996 and 2002, respectively. He is currently a faculty member in the Department of Information Systems, City University of Hong Kong. His research interests include image compression, motion estimation/compensation, e-Learning, and information retrieval. He is a member of the IEEE.



**Jin-Xing Hao** is a PhD candidate in the Department of Information Systems, City University of Hong Kong, Kowloon, Hong Kong. His research interests include both the technical and the behavioral aspects of information systems such as knowledge engineering, social networking, e-learning, and organizational theories.

► **For more information on this or any other computing topic, please visit our Digital Library at [www.computer.org/publications/dlib](http://www.computer.org/publications/dlib).**